Challenging the Manifesto Project data monopoly: Estimating policy position time-series using expert and mass survey data

Bastiaan Bruinsma[1] and Kostas Gemenis[2]

[1]Scuola Normale Superiore

[2]Universiteit Twente

Paper prepared for presentation at the 2017 general conference of the European Political Science Association, Milan, 22-24 June, 2017

Abstract

Whenever researchers need to test theories and hypotheses using longitudinal data of political parties' ideological and policy placement, they have little choice. Researchers are often constrained to use the Manifesto Project data, despite the extensive evidence that has challenged its reliability and validity. In this paper we show that it is possible to construct a unique and rich time-series of policy placements by combining expert and mass survey data, and addressing the problem of missing values through the Amelia II multiple imputation algorithm. Using data from Germany, the Netherlands, and Greece, we estimate the positions of parties on the left-right dimension and on a two-dimensional (socio-economic and socio-cultural) space, and show how the estimates outperform the Manifesto Project estimates in terms of their face validity.

1 Introduction

Scholars interested in longitudinal data on political parties' ideological decisions face a significant problem early on in their studies: the lack of substantive data sets. While comparable cross-national party positions are available in a plenitude of forms, the same cannot be said for longitudinal party positions. In most cases, researchers have little choice but to turn to the data generated by the Manifesto Project, despite the substantive evidence that has challenged the reliability and validity of the project (Gemenis 2013a). This is rather unfortunate, since using data with poor measurement qualities, is likely to lead to erroneous inferences with regard to the theories and hypotheses political scientists and other social scientists put to test. Put simply, our empirical findings about many of these theories and hypotheses may well be wrong because of the poor quality of available data, as has been repeatedly demonstrated for theories of voting (Lewis and King 1999, pp. 26-31), government coalition formation (Strøm, Müller, and Bergman 2003, pp. 8-9), party competition (Benoit, Laver, and Mikhaylov 2009, pp. 507-510), and citizens' policy representation (Golder and Stramski 2010, pp. 98-99) just to name a few.

In addition, having only a single, and imperfect, longitudinal source of party positions makes it difficult to cross-validate new methods for longitudinal party placement such as those arising from automated text analyses (e.g. Laver, Benoit, and Garry 2003; Slapin and Proksch 2008). As a result, the proponents of estimating policy positions by means of automated text analysis have long abandoned the use of the Manifesto Project data as a gold standard (Benoit and Laver 2007a) in favour of expert survey estimates. However, with the exception of the Chapel Hill Expert Survey trend file that goes back to 1999 (Bakker et al. 2015), expert surveys have been conducted in a rather haphazard manner.

With this paper we aim to challenge the counter-productive monopoly of the Manifesto Project by combining all the available expert and mass survey estimates of parties' policy positions into a single dataset and work out through the missing data problem by means of multiple imputation. We do so we use the **Amelia** package for R (Honaker, King, and Blackwell 2011) which allows us to flexibly impute the data and insert different time effects. To build the data-set we will draw on Eurobarometer, CSES and CHES data, as well as several other comparative projects and country specific surveys.

The paper will run as follows. First, we will turn briefly to the methodological problems relating to validity and reliability that have been raised with the Manifesto Project, and discuss the advantages and disadvantages of employing expert and mass survey data to measure parties' policy positions. Subsequently, we explain the logic of multiple imputation and introduce the **Amelia** package. We then turn to our cases and use **Amelia** to generate data-sets for the Netherlands, Germany, and Greece and discuss the results in comparison to the respective estimates given by the Manifesto Project.

2 Estimating parties' policy positions

Without a doubt, the longest established and most popular approach to estimate parties' policy positions has involved the hand-coding of party manifestos. The most important data generation project associated with the method is the Manifesto Project which has been running uninterrupted since 1979.[1] The Manifesto Project approach to estimation involves the selection of a document that is considered to function as a manifesto for a particular election, and the coding of each and every 'quasi-sentence' of the document into a set of predefined categories. The resulting data are then scaled to ideological dimensions of interest, with the most famous being the general left-right dimension ('rile' in the Manifesto Project datasets), that has been flagged by the project as its 'crowning achievement' (Budge et al. 2001, p.19).

Even though the project has been quite popular since its inception (as measured in terms of usage and citations), it has received a substantive amount of criticism during the past two decades. It has been argued that the selected documents are not always comparable in terms of their policy coverage (Gemenis 2012; Hansen 2008) which has implications in terms of the quality of the produced estimates. Moreover, it has been shown that the hand-coding process used by the Manifesto Project is notoriously unreliable (Mikhaylov, Laver, and Benoit 2012), while most of the documents in its datasets seem to have been coded by coders who performed poorly on the coder reliability training test (Gemenis 2013b, pp. 9-12).[2]

The original left-right ('rile') scale proposed by the Manifesto Project has been the

[1]Formerly known as the Manifesto Research Group (MRG, 1979–1989), the Comparative Manifesto Project (CMP, 1989–2009), and Manifesto Research on Political Representation (MARPOR, 2009–).

[2]Traditionally, the Manifesto Project has responded to coding reliability criticisms by arguing that the reliability of the date is generally higher than measured due to the training of coders (e.g. Volkens, Bara, and Budge 2009), but the data from the coder training test provided by the Project itself presented in Gemenis (2013b) points otherwise. Moreover, the Project's own assessment of inter-coder reliability (Lacewell and Werner 2013) largely confirmed the assessment made by Mikhaylov, Laver, and Benoit (2012). Consequently, Lacewell and Werner (2013) have tried to downplay the small size of the estimated reliability coefficients.

subject of criticism in a number of different published journal articles (e.g. Benoit and Laver 2007b; Dinas and Gemenis 2010; Gabel and J. D. Huber 2000; Pelizzo 2003) in terms of flawed assumptions in the construction of the left-right index and/or lack of validity of the produced estimates. While, most of the critiques have tried to scale the Manifesto Project data in different ways in order to improve the validity of the resulting estimates (with some interesting debates among the critics such as those between Franzmann (2015), D. Jahn (2011), and Detlef Jahn (2014)), it begs the question whether there is any point in searching for the best scaling model given the fundamentally flawed and unreliable nature of the input data.

Therefore, researchers have started to turn to other methods of establishing party positions such as expert surveys or mass surveys in their empirical applications (e.g. Bäck and Dumont 2007; Golder and Stramski 2010; Vowles and Xezonakis 2016; Warwick 2006). They do so not only for illustrative purposes or hypothesis testing, but also use mostly expert surveys as a benchmark in many 'proof-of-concept' illustration of new methods (such as Wordscores (Laver, Benoit, and Garry 2003), Wordfish (Slapin and Proksch 2008), Delphi coding (Gemenis 2014), and crowd-sourcing (Benoit, Conway, et al. 2016)). Instead of relying on coders, expert studies derive their estimates from experts, who provide parties' positions directly on the ideological dimensions of interest, with the mean mean judgment across the experts as the party position estimate. A similar procedure is followed for mass surveys, but instead of a select group of experts, large groups of citizens are asked to provide estimates of parties' positions directly on ideological dimensions of interest, with again their mean taken as the party position estimate. Both approaches are therefore flexible and provide immediately usable information regarding the parties' positions.

Nevertheless, expert and mass surveys have problems of their own. Especially for expert surveys it is hard to establish parties' positions retrospectively unless we accept large measurement error due to telescoping effects, whereby experts recall the facts for the wrong time period (Steenbergen and Marks 2007, p.349). In addition, especially in the case of new or small parties, experts might disagree on how to place these parties resulting in unreliable estimates. And while proponents of expert surveys have acknowledged and tried to explain this disagreement (Steenbergen and Marks 2007), they have not yet proposed how to reduce it. Similar problems arise with voters who, when having little information about politics are known to provide erroneous judgments of party placements (Tilley and Wlezien 2008). Moreover, cognitive dissonance leads to projection biases, where citizens do not judge parties objectively but in relation to their own position and sympathy for them (Merrill, Grofman, and Adams 2001). Experts' are not resistant to this threat either as it has been shown that their own ideological preferences introduce bias to estimates for certain type of parties (Curini 2010). Finally, most expert surveys suffer from limited data availability being either one-off events or having been around only since quite recently. The Chapel Hill Expert Survey (CHES) has been ongoing only since 2002 (Hooghe et al. 2010) and does not retrospectively generate data before that time, while the Comparative Study of Electoral Systems (CSES) is ongoing since 1996 but only in countries where they have resources to field a national election study.

Here, we try to built on the strengths of expert and mass surveys while at the same time addressing their shortcomings. By not using retroactive studies, we prevent effects such as telescoping, while using multiple imputation allows us to generate long time series. The next section will discuss how this imputation procedure works, after which we will demonstrate its use in the case of the Netherlands, Germany and Greece.

3 Multiple Imputation

Multiple imputation (MI) is based around the idea that missing values are not imputed only once, but multiple times. The different imputed values this generates represent the degree of uncertainty we have regarding what the imputed value should be. When there is high number of missing values and thus high uncertainty, the difference between the imputed values is large, while in the opposite case the difference between the values is low (King et al. 2001). As such, MI improves on other imputation procedures such as mean- or last value carried forward imputation by not only giving an estimate of the missing value, but also an estimate of how sure we are of that value (Horton and Kleinman 2007).

In order to run successfully, MI assumes that the data it uses is *missing at random* (MAR) (see also Table 1. Despite the wording, this does not mean that the missingness is random, but rather that the missingness is only related to the observed values and not to non-observed ones. We can therefore predict the missing values with the information contained in the observed values. What most would understand as random missingness is

Acronym	Meaning	Predication possible with:
MCAR	Missing Completely at Random	-
MAR	Missing at Random	Non-missing data
MNAR	Missing Not at Random	Non-missing and missing data

Adapted from King et al. (2001)

Table 1: The three types of missingness

instead labelled as data that is *missing completely at random* (MCAR). Here, we cannot predict missing values by using either the observed or the non-observed data. An example of MCAR would occur when respondents will throw a die before answering any question and only respond when it shows 6. Of course, such things rarely happen and the MCAR assumption therefore rarely applies. The third variant, data that is *missing not at random* (MNAR) or is *non-ignorable* (NI) is the most problematic. Here, the missingness depends not only on the observed data, but on the missing data as well. We therefore have no way of knowing why the data is missing and what its eventual value should be.

To determine whether or not the data is MCAR or MAR, one can either use Little's test (Little 1988) or create dummy variables for missingness and see whether the missingness of one variable relates to any of the other variables. However, in most real-life applications the distinction will mostly be between MAR or MNAR. As there is no direct test for this, categorising data as one of the two depends on the argumentation of the researcher. In our case, we deem the assumption of MAR appropriate as the reason the data is missing is known (there were no expert surveys that year) and is not dependent on the missing data itself (the reason a party's position is missing for a year is not because of the position of the party).

While here we advocate using Amelia for purposes of multiple imputation, we want to stress that other methods for imputation are also available. Within R, there are packages as **missForest**, **mice** and **mi** that offer viable alternatives. Each of these packages differs mainly in the options they offer and the method they choose for implementing (multiple) imputation. While the main distinction in the methods is still that between *single* and *multiple* imputation, most methods focus upon the latter as it allows for calculation of error around the imputed value.

- Paragraph on machine learning - Robustness check with the other MI methods for CDU/CSU

4 Amelia

The **Amelia** package for R (Honaker, King, and Blackwell 2011) implements MI using a bootstrapping-based Expectation-Maximization algorithm (Horton and Kleinman 2007). As such, the method is able to generate fast and independent imputations even with small samples and a large number of parameters (King et al. 2001). Apart from assuming MAR, another assumption **Amelia** makes is that the complete set of observed and unobserved data is multivariate normal. When we denote the dataset itself as D (with rows and columns $n \times k$), then the multivariate normal assumption is:

$$D \sim \mathcal{N}_k(\mu, \Sigma) \tag{1}$$

showing that D has a multivariate normal distribution with mean μ and co-variance matrix Σ. While the multivariate assumption is almost never entirely correct, Honaker and King (2010) suggest that ample evidence suggests that the model works as well as more complicated models. For the MAR assumption, if we have M denote the matrix that indicates whether or not data is missing, this assumption is:

$$p(M|D) \;=\; p(M|D^{obs}) \tag{2}$$

This means that the pattern of missingness depends on the observed, and not on the missing data. To impute the missing values, we need to gain an understanding of what the complete-data looks like. This complete data has the parameters $\theta = (\mu, \Sigma)$. As the observed data (that is, the data we actually know) is the observed data itself (D^{obs}) and our knowledge of the missing values (M), the likelihood of our data is $p(D^{obs}, M|\theta)$. Using the MAR assumption, this becomes:

$$p(D^{obs}, M|\theta) \;=\; p(M|D^{obs})p(D^{obs}|\theta) \tag{3}$$

As we carry out the inference on the complete data parameters, the likelihood becomes

$$L(\theta|D^{obs}) \propto p(D^{obs}|\theta) \tag{4}$$

which can be rewritten as

$$p(D^{obs}|\theta) \ = \ \int p(D|\theta)dD^{mis} \tag{5}$$

Giving us, with a flat prior on θ, the following posterior:

$$p(\theta|D^{obs}) \propto p(D^{obs}|\theta) \ = \ \int p(D|\theta)dD^{mis} \tag{6}$$

The EMB algorithm (Honaker and King 2010) then bootstraps (B) the data to simulate the estimation uncertainty and subsequently runs the expectation-maximization (EM) algorithm to find the mode of the posterior for the bootstrapped data. Then, Amelia draws values of D^{mis} from the distribution based on the complete-data parameters, which are conditional on D^{obs} and the draws of θ. This generates a specific number of imputed data sets, which can be analyzed separately using similar methods as those that assume complete-data. The results of these separate analyses can then be combined using the rules set out by Rubin (1987), or using the **Zelig** package for R (Choirat et al. 2017; Imai, King, and Lau 2008).

5 Case selection, Sources and Methods

Regarding the number of variables to include in the imputation model, Rubin (1996) advises to retain as many variables as possible. In that line, Honaker and King (2010) advise to also include the dependent variables (the Manifesto Project estimates in our case), as imputation models are predictive, and not causal, models. Nevertheless, as here our goal is not so much as to prove a relation between the dependent and the independent variables as it is to build a new data-set, we will not include the dependent variables. In addition, while the common advice is that 5-10 imputations are adequate (Rubin 1987), we follow Bodner (2008) and take as many imputations as the percentage of missing data in the respective data set. We do so because of the high missingness among some of the variables and because running this number of imputations does not cause major problems on any modern computer.

We choose three countries with different characteristics to see whether imputation can generate use-full data-sets: the Netherlands, Germany, and Greece. The Netherlands has a considerable number of expert surveys over the full period from 1945 onward.

Germany has data of similar quality, but only from 1984. Greece starts somewhat earlier in 1980, but has only a few expert surveys of varying quality. In addition, while Germany experienced a substantial change during the reunification in 1990 and Greece had a new party system after 1974, the party system in the Netherlands remained relatively stable.

To build the datasets for these countries we draw on expert surveys and mass surveys. The expert surveys can be either singular, or may have occurred on multiple occasions such as the Chapel Hill Expert Survey (CHES) (Bakker et al. 2015; Polk et al. 2017), which has been used in each of the three countries. Also, some expert surveys (e.g. Benoit and Laver 2006; Castles and Mair 1984; Kitschelt 2011) were available for all countries, while others (e.g. Gemenis and Nezi 2012; Janda 1980) were only available for a single country. For the mass surveys we used Eurobarometer data for all three the countries and data from the Dutch Parliamentary Election Studies (DPES) for the Netherlands, the German Longitudinal Election Study (GLES) and Politbarometer for Germany, data collected by EKKE (National Centre for Social Research) in Greece. For a full overview of the sources utilized for each country we refer to Appendix A.

For each of the countries, we use the expert surveys to create a standardized variable *Expert L-R* which has the left-right positions for each party based on the expert surveys for each of the years that we could find in our data collection, and missing values for the years in which no surveys could be found. In addition, we created an *Expert EC* and *Expert SO* variable in a similar way to help us better estimate the missing data for the *Expert L-R* variable. For similar reasons, we add variables for the left-right and materialist-post-materialist positions based on Eurobarometer data, and variables for left-right positions based on national election data.

For each of the three countries we then run **Amelia** with the different parties set as the cross-section variable and years as the time-series variable. In addition, we include a cubic polynomial to the imputation model to account for the effects in time and make this polynomial vary across the different parties. These steps allow **Amelia** to distinguish the different parties over time and separate the different time effects for each of them. Finally, we set a ridge prior which helps us to deal with the problems caused by the high missingness in the data, which for each of the three countries can be as high as 80% for some of the variables. The ridge prior does so by diminishing the co-variances of the data while keeping the means and variances the same. Following the advice in Honaker, King, and Blackwell (2011), we set this ridge prior at 1%. Finally, because of the high missingsness and following the advice of Bodner (2008) we set the number of imputations at 80. Once we have run **Amelia** with these settings we combine the imputations and calculate the mean for each of the variables which we then plot together with a loess regression line to observe the variation over time. In addition, we plot the positions given by 'rile' using the **manifestoR** package for R which allows direct integration of Manifesto Project data into R (Lehmann et al. 2016; Volkens, Lehmann, et al. 2016). We use the same package to derive the codings that allow us to construct two new measures for *economic* and *social* policy based on the indexes suggested by Benoit and Laver (2007b).

6 Results

Here we show the results for the left-right dimension for each of the three countries. For each of the graphs with the imputed data, the positions run from 0 (Left) to 1 (Right), while for RILE they run from -50 (Left) to 50 (Right). In the graph with the imputed data, the dots indicate the mean imputation for that particular party in a specific year, with a loess regression line based on those points running through them. For the graph with the 'rile' data the dots represent the values as given by the Manifesto project which we simply connected. Finally, note that not all parties that were included in the imputation model are shown here to prevent the graphs from becoming too cluttered. Moreover, the different countries have different time spans - the Netherlands ranges from 1945 to 2012, Germany from 1984 to 2013, and Greece from 1974 to 2015. Here we show only the results for 'rile' but similar results for the *economic* and *social* policy dimensions can be found in Appendix B.

Figures 1 and 2 show the results for the Netherlands, with 6 parties shown. In the graph based on the imputed data, the positions of the parties are shown to be stable over time and to fluctuate only minimally. Each of the parties basically keeps its ordinal position and does not cross any of the other parties. This in contrast with the graph based on the 'rile' data, which fluctuates heavily and were parties cross each other on multiple occasions. In that respect, the expert positions and the imputed values are less extreme with only a few of the imputed values located far from the loess line.

Figures 3 and 4 show similar results for Germany, with the CDU/CSU, FDP, B'90/ Die

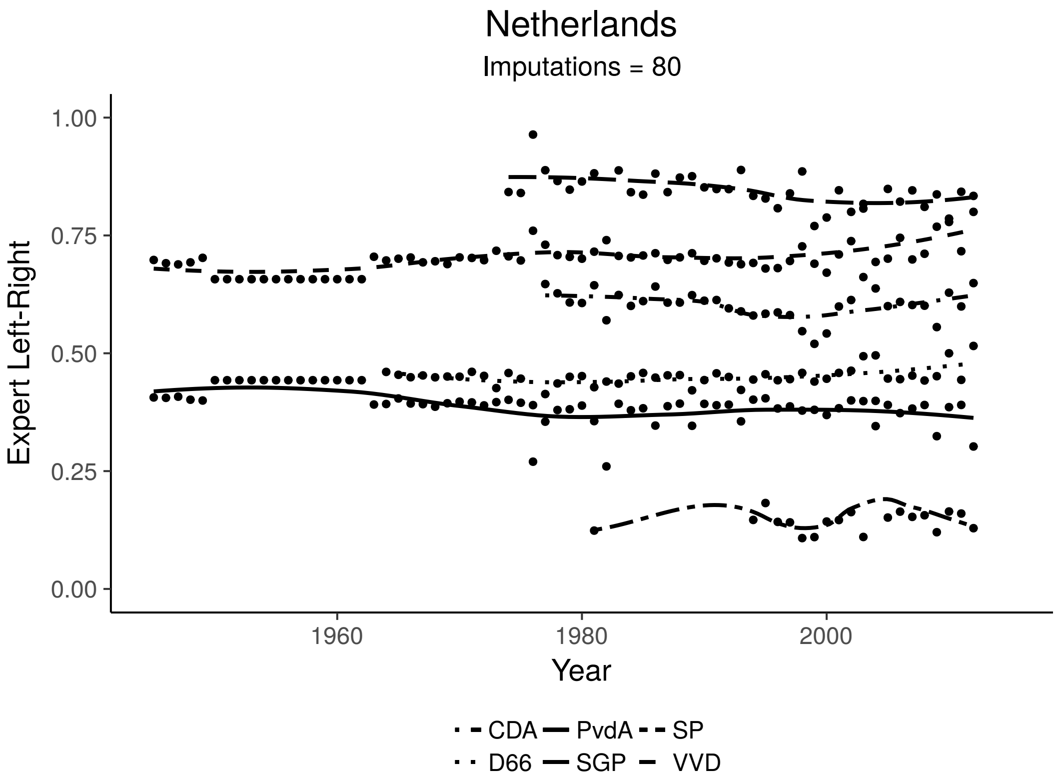


Figure 1: Imputed Expert Positions for the Netherlands

Gr unen, SPD, PDS and DIE LINKE shown. As with the Netherlands, a similar pattern of stability is found in the imputed data which misses in the RILE data. The order of the parties with the CDU/CSU being the most right-wing and PDS/DIE LINKE being the most left-wing remains stable over time and also the ordering of the other parties seems correct. This in contrast to the RILE data in which the SPD and CDU/CSU sometimes switch position. Please note that the disruption for the most left party in the graph is caused by the PDS and DIE LINKE being treated as two separate parties in the imputation model.

Finally, figures 5 and 6 show the results for Greece, with the KKE, ND, PASOK, Synaspismos and SYRIZA shown. Here, the difference between the two graphs is most obvious, with the RILE data actively crossing each other, while the imputed data remains more stable. Notice for example the KKE data generated by 'rile' which fluctuates heavily over time and at certain points is even positioned as the most right-wing party, while in the imputed data it is always the most left-wing party. A similar story goes for the other parties whose position (right wing for the ND and mid-left to extreme left positions for all the other parties) can clearly be distinguished in the imputed data, but not in the

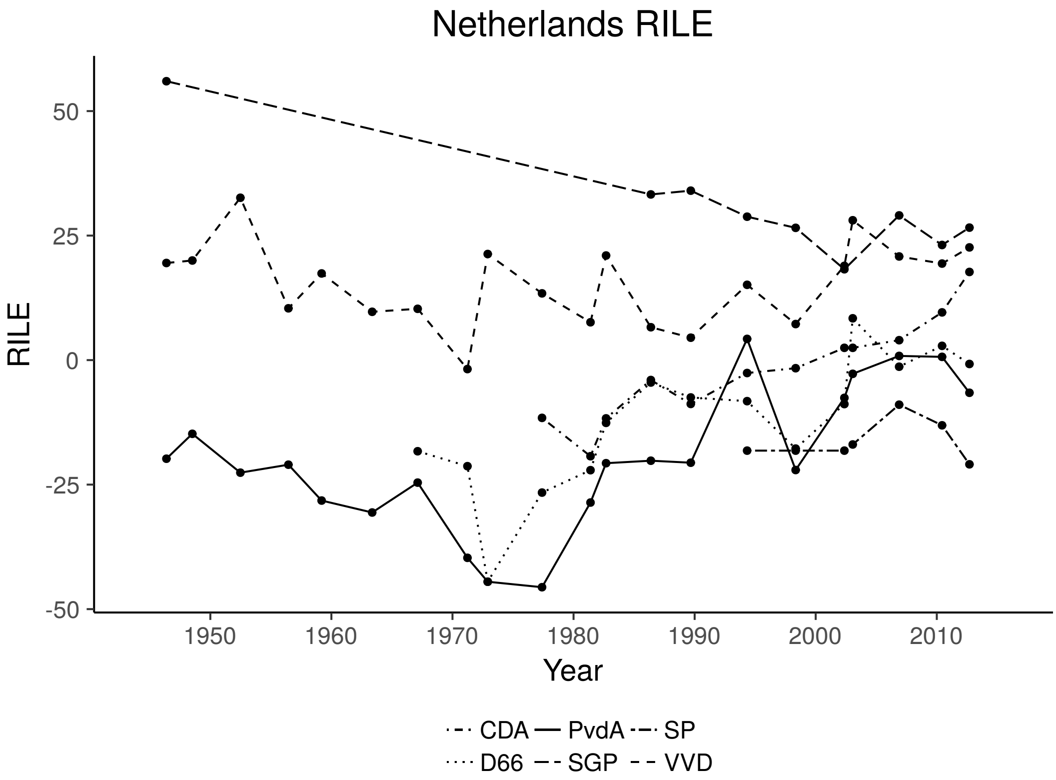


Figure 2: RILE Positions for the Netherlands

'rile' data.

7 Discussion

Our aim in this paper was to challenge the Manifesto Project monopoly by providing an alternative data-set which we built by combining all the available expert and mass survey estimates of parties' policy positions into a one dataset and addressing the missing data problem by means of multiple imputation. Solely based on face validity, we have shown that such a data-set imputation technique can be successfully used to compute relatively stable party positions over time. This despite the different sources of the positions and the often high missingness. Moreover, we found that the procedure worked similarly well in countries where there is a considerable amount of high quality data such as the Netherlands and Germany, as for countries in which very few and often questionable data is available such as Greece. This indicates that similar results should be able to achieve in other countries, though we may require a broader case selection to validate this point.

We belief the greatest contribution of this work lies in establishing a novel way in

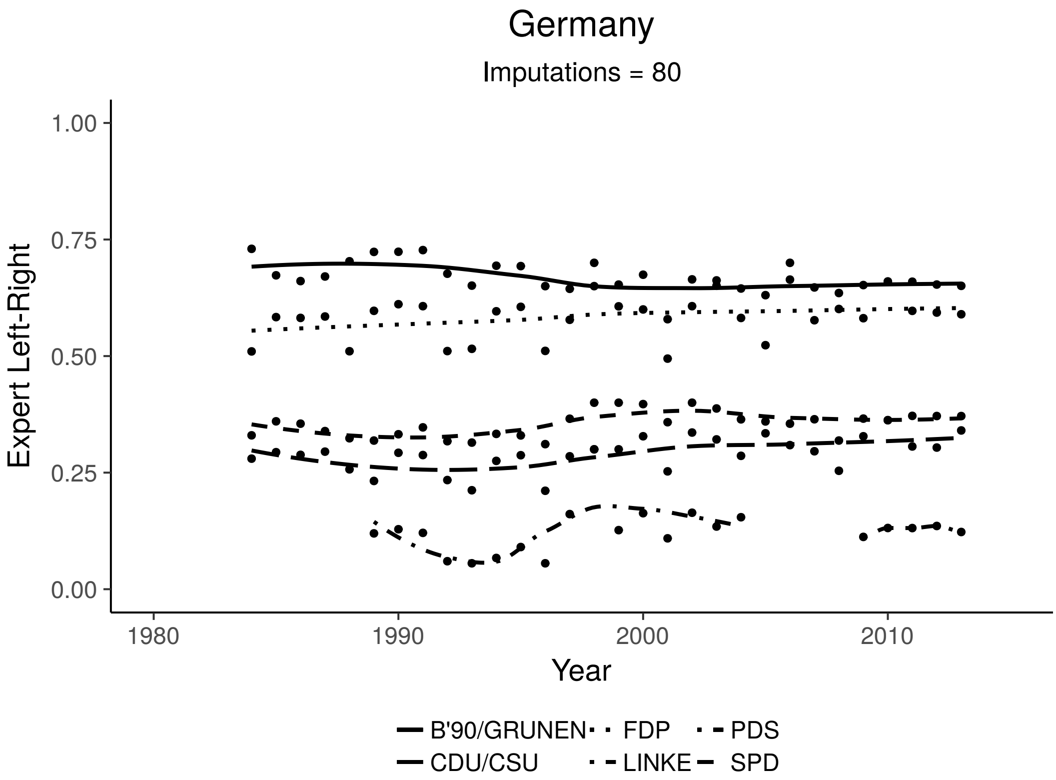


Figure 3: Imputed Expert Positions for Germany

which methods of automated content analysis might be validated. Right now they have often been compared to Manifesto Project data - which is often questioned - or have been proven by face validity alone. We belief that comparing any estimates against an imputed time-series of expert survey data might prove to be a good alternative. In addition, the multiple imputation method used here, Amelia, is flexible and allows for even more features than we have discussed here. For example, the method allows for the integration of Bayesian priors, splines and lags and leads that could improve the imputation in certain circumstances. This will allow scholars to be as flexible as need be and tailor the imputation to the specifics of the country. In the end, this will lead to a rich data-set that we argue more accurately captures the positions of countries over time than the values generated by the Manifesto Project data.

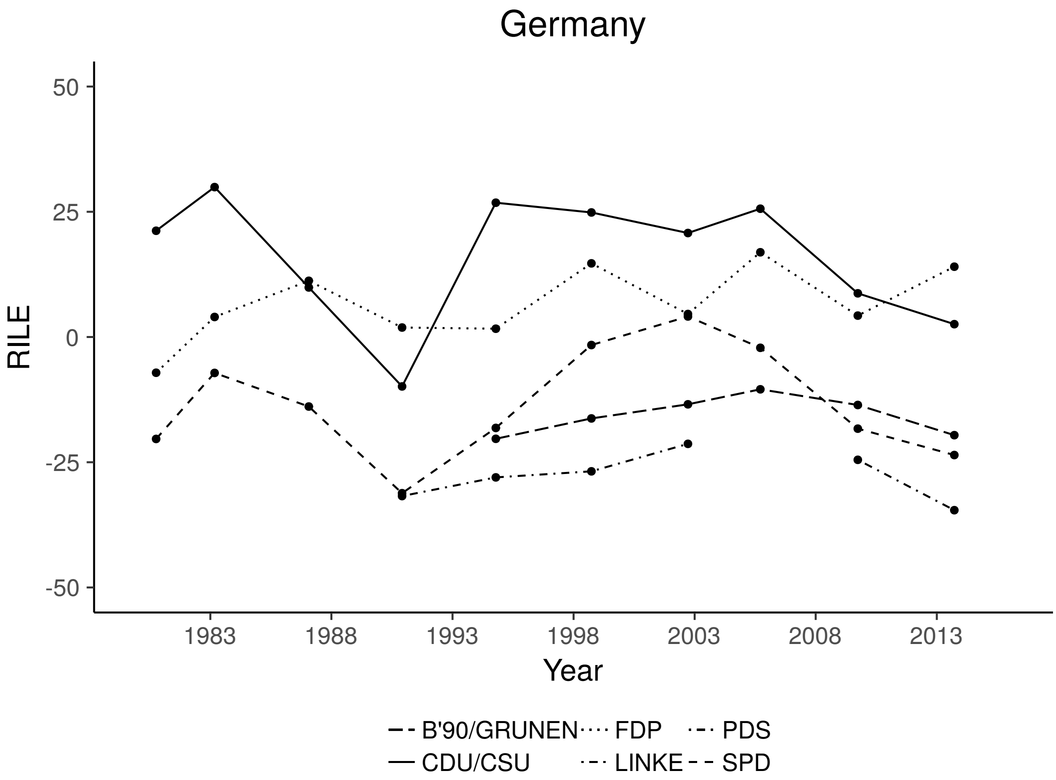


Figure 4: RILE Positions for Germany

References

Bäck, Hanna and Patrick Dumont (2007). "Combining large-n and small-n strategies: The way forward in coalition research". In: *West European Politics* 30.3, pp. 467–501. eprint: `http://dx.doi.org/10.1080/01402380701276295`.

Bakker, Ryan et al. (2015). *2014 Chapel Hill Expert Survey*. Version 2015.1. University of North Carolina, Chapel Hill. URL: `http://www.chesdata.eu`.

Benoit, Kenneth, Drew Conway, et al. (2016). "Crowd-sourced Text Analysis: Reproducible and Agile Production of Political Data". In: *American Political Science Review* 110.2, pp. 278–295.

Benoit, Kenneth and Michael Laver (2006). *Party Policy in Modern Democracies*. London: Routledge.

— (2007a). "Benchmarks for text analysis: A response to Budge and Pennings". In: *Electoral Studies* 26.1, pp. 130–135.

— (2007b). "Estimating party policy positions: Comparing expert surveys and hand-coded content analysis". In: *Electoral Studies* 26.1, pp. 90–107. ISSN: 0261-3794.

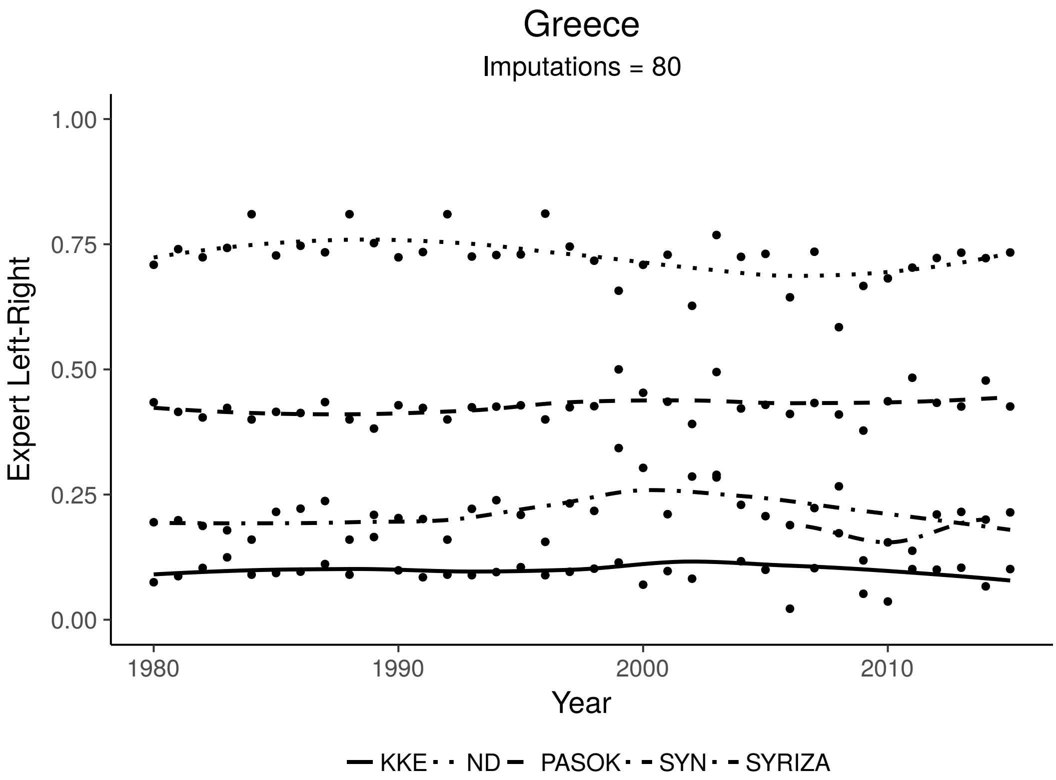


Figure 5: Imputed Expert Positions

Benoit, Kenneth, Michael Laver, and Slava Mikhaylov (2009). "Treating words as data with error: Uncertainty in text statements of policy positions". In: *American Journal of Political Science* 53.2, pp. 495–513.

Bodner, Todd E. (2008). "What Improves with Increased Missing Data Imputations?" In: *Structural Equation Modeling: A Multidisciplinary Journal* 15.4, pp. 651–675. eprint: `http://dx.doi.org/10.1080/10705510802339072`.

Budge, Ian et al. (2001). *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945-1998.* Oxford, UK: Oxford University Press. ISBN: 0-19-924400-6.

Castles, Francis G. and Peter Mair (1984). "Left-Right Political Scales: Some 'Expert' Judgments". In: *European Journal of Political Research* 12.1, pp. 73–88.

Choirat, Christine et al. (2017). *Zelig: Everyone's Statistical Software.* Version 5.1-1.

Curini, Luigi (2010). "Experts' Political Preferences and their Impact on Ideological Bias - An Unfolding Analysis based on a Benoit-Laver Expert Survey". In: *Party Politics* 16.3, pp. 299–321.

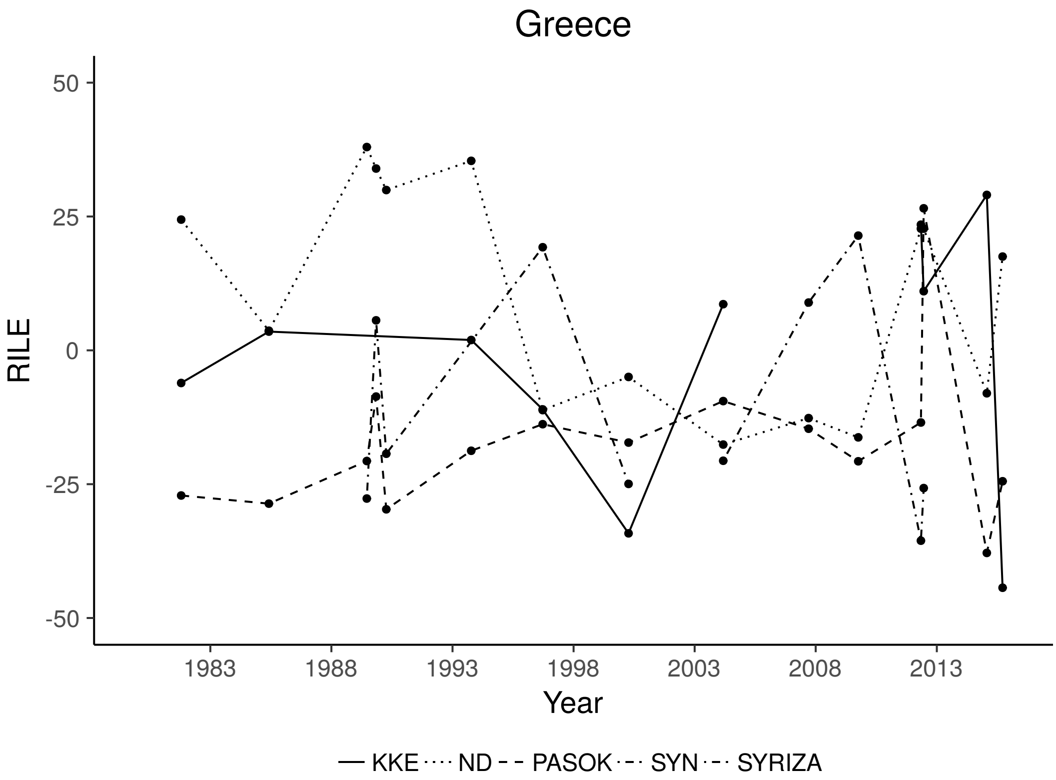


Figure 6: RILE Positions

Dinas, Elias and Kostas Gemenis (2010). "Measuring Parties' Ideological Positions With Manifesto Data: A Critical Evaluation of the Competing Methods". In: *Party Politics* 16.4, pp. 427–450. ISSN: 1354-0688.

Franzmann, Simon T. (2015). "Towards a real comparison of left-right indices". In: *Party Politics* 21.5, pp. 821–828. eprint: `http://dx.doi.org/10.1177/1354068813499865`.

Gabel, Matthew J. and John D. Huber (2000). "Putting Parties in Their Place: Inferring Party Left-Right Ideological Positions from Party Manifestos Data". In: *American Journal of Political Science* 44.1, pp. 94–103.

Gemenis, Kostas (2012). "Proxy documents as a source of measurement error in the Comparative Manifestos Project". In: *Electoral Studies* 31, pp. 594–604.

— (2013a). "Estimating parties' policy positions through voting advice applications: Some methodological considerations". In: *Acta Politica* 48.3, pp. 268–295.

— (2013b). "What to Do (and Not to Do) with the Comparative Manifestos Project Data". In: *Political Studies* 61, pp. 3–23. ISSN: 1467-9248.

— (2014). "An iterative expert survey approach for estimating parties' policy positions". In: *Quality & Quantity* 49.6, pp. 2291–2306. ISSN: 15737845.

Gemenis, Kostas and Roula Nezi (2012). *The 2011 political parties expert survey in Greece.* Data Archiving and Networked Services (DANS) [distributor].

Gemenis, Kostas and Carolien van Ham (2014). "Comparing methods for estimating parties ' positions in Voting Advice Applications". In: *Matching Voters with Parties and Candidates: Voting Advice Applications in a Comparative Perspective*. Ed. by Diego Garzia and Stefan Marschall. Colchester, UK: ECPR Press. Chap. 3, pp. 33–47. ISBN: 9781907301735.

Golder, Matt and Jacek Stramski (2010). "Ideological Congruence and Electoral Institutions". In: *American Journal of Political Science* 54.1, pp. 90–106. ISSN: 1540-5907.

Hansen, Martin Ejnar (2008). "Back to the archives? A critique of the Danish part of the manifesto dataset". In: *Scandinavian Political Studies* 31.2, pp. 201–216.

Hix, Simon and Christopher Lord (1997). *Political Parties in the European Union*. New York, NY: St. Martin's Press. ISBN: 978-0-333-60921-7.

Honaker, James and Gary King (2010). "What to do About Missing Values in Time Series Cross-Section Data". In: *American Journal of Political Science* 54.2, pp. 561–581.

Honaker, James, Gary King, and Matthew Blackwell (2011). "Amelia II: A Program for Missing Data". In: *Journal of Statistical Software* 45 (7), pp. 1–47.

Hooghe, Liesbet et al. (2010). "Reliability and validity of the 2002 and 2006 Chapel Hill expert surveys on party positioning". In: *European Journal of Political Research* 49.5, pp. 687–703.

Horton, Nicholas J. and Ken P. Kleinman (2007). "Much Ado About Nothing: A Comparison of Missing Data Methods and Software to Fit Incomplete Data Regression Models". In: *The American Statistician* 61.1, pp. 79–90.

Huber, John and Ronald Inglehart (1995). "Expert Interpretations of Party Space and Party Locations in 42 Societies". In: *Party Politics* 1.1, pp. 73–111.

Imai, Kosuke, Gary King, and Olivia Lau (2008). "Toward A Common Framework for Statistical Analysis and Development". In: *Journal of Computational Graphics and Statistics* 17.4, pp. 892–913.

Jahn, D. (2011). "Conceptualizing Left and Right in comparative politics: Towards a deductive approach". In: *Party Politics* 17.6, pp. 745–765. ISSN: 1354-0688.

Jahn, Detlef (2014). “What is left and right in comparative politics? A response to Simon Franzmann”. In: *Party Politics* 20.2, pp. 297–301. eprint: http://dx.doi.org/10.1177/1354068813520273.

Janda, K. (1980). *Political Parties: A Cross-National Survey.* New York, NY: Free Press.

King, Gary et al. (2001). “Analyzing Incomplete Political Science Data: An Alternative Algorithm for Multiple Imputation”. In: *American Political Science Review* 95.1, pp. 49–69.

Kitschelt, Herbert (2011). *Democratic Accountability and Linkages Project.* Durham, NC: Duke University.

Lacewell, O. P. and A. Werner (2013). “Divided We Fall? Polarization in the 2012 U.S. Presidential Election”. In: *Die US-Präsidentschaftswahl 2012 Analysen der Politik- und Kommunikationswissenschaft.* Ed. by C. Bieber & K. Kamps. Wiesbaden: Springer Fachmedien Wiesbaden GmbH.

Laver, Michael, Kenneth Benoit, and John Garry (2003). “Extracting Policy Positions from Political Texts Using Words as Data”. In: *The American Political Science Review* 97.2, pp. 311–331.

Laver, Michael and W. Ben Hunt (1992). *Policy and Party Competition.* New York, NY: Routledge.

Laver, Michael and Peter Mair (1999). “Party Policy and Cabinet Portfolios in the Netherlands 1998: Results from an Expert Survey”. In: *Acta Politica* 34.1, pp. 49–66.

Lehmann, Pola et al. (2016). *Manifesto Corpus.* Version 2016-6. Berlin: WZB Berlin Social Science Center.

Lewis, Jeffrey and Gary King (1999). “No Evidence on Directional vs. Proximity Voting”. In: *Political Analysis* 8 (1), pp. 21–33.

Little, Roderick J. A. (1988). “A Test of Missing Completely at Random for Multivariate Data with Missing Values”. In: *Journal of the American Statistical Association* 83.404, pp. 1198–1202.

Lubbers, Marcel (2000). *Expert Judgment Survey of Western-European Political Parties 2000.* Nijmegen, NL: NWO, Afdeling Sociologie - Universiteit Nijmegen.

Merrill, Samuel, Bernard Grofman, and James Adams (2001). “Assimilation and contrast effects in voter projections of party locations: Evidence from Norway, France, and the USA”. In: *European Journal of Political Research* 40.2, pp. 199–221. ISSN: 03044130.

Mikhaylov, Slava, Michael Laver, and Kenneth Benoit (2012). "Coder Reliability and Misclassification in the Human Coding of Party Manifestos". In: *Political Analysis* 20.1, pp. 78–91. eprint: http://pan.oxfordjournals.org/content/20/1/78.full.pdf+html.

Morgan, Michael John (1976). "The Modelling of Government Coaition Formations: A Policy-Based Approach with Interval Measurement". PhD thesis. Ann Arbor, MI: University of Michigan.

Pelizzo, Riccardo (2003). "Party Positions or Party Direction? An analysis of Party Manifesto Data". In: *West European Politics* 26.2, pp. 67–89.

Polk, Jonathan et al. (2017). "Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill Expert Survey data". In: *Research & Politics* 4 (1), pp. 1–9. eprint: http://dx.doi.org/10.1177/2053168016686915.

Ray, Leonard (1999). "Measuring party orientations towards European integration: results from an expert survey". In: *European Journal of Political Research* 36.2, pp. 283–306. ISSN: 1475-6765.

Rubin, Donald B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York, NY: John Wiley & Sons.

— (1996). "Multiple Imputation After 18+ Years". In: *Journal of the American Statistical Association* 91.434, pp. 473–489.

Schmitt, H. et al. (2008). *The Mannheim Eurobarometer Trend File 1970-2002*. Data file. Version 2. European Commission [Principal investigator].

Slapin, Jonathan B. and Sven-Oliver Proksch (2008). "A Scaling Model for Estimating Time-Series Party Positions from Texts". In: *American Journal of Political Science* 52.3, pp. 705–722. ISSN: 0092-5853.

Steenbergen, Marco R. and Gary Marks (2007). "Evaluating expert judgments". In: *European Journal of Political Research* 46, pp. 347–366.

Strøm, Kaare, Wolfgang C. Müller, and Torbjörn Bergman (2003). *Delegation and Accountability in Parliamentary Democracies*. Oxford, UK: Oxford University Press.

Tilley, James and Christopher Wlezien (2008). "Does Political Information Matter? An Experimental Test Relating to Party Positions on Europe". In: *Political Studies* 56.1, pp. 192–214.

Volkens, Andrea, Judith Bara, and Ian Budge (2009). "Data Quality in Content Analysis. The Case of the Comparative Manifesto Project". In: *Historical Social Research* 34.1, pp. 234–251.

Volkens, Andrea, Pola Lehmann, et al. (2016). *The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR)*. Version Version 2016b. Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).

Vowles, Jack and Georgios Xezonakis (2016). *Globalization and Domestic Politics*. Oxford, UK: Oxford University Press.

Warwick, Paul V. (2006). *Policy Horizons and Parliamentary Government*. Basingstoke, UK: Palgrave Macmillan.

Appendix A - Data Sources

Expert	Type	Period
Janda (1980)	Expert	1950-1962
Morgan (1976)	Expert	1976
Castles and Mair (1984)	Expert	1982
Laver and Hunt (1992)	Expert	1989
J. Huber and Inglehart (1995)	Expert	1993
Laver and Mair (1999)	Expert	1998
Lubbers (2000)	Expert	2000
Benoit and Laver (2006)	Expert	2003
Kitschelt (2011)	Expert	2009
Gemenis and van Ham (2014)	Expert	2012
Chapel Hill Expert Survey (2015; 2017)	Expert	1999, 2002, 2006, 2010
Eurobarometer	Survey	1973, 1976-1997, 2000, 2008
Dutch Parliamentary Election Survey	Survey	1977-2012

Table 2: Overview of the data employed for the Netherlands

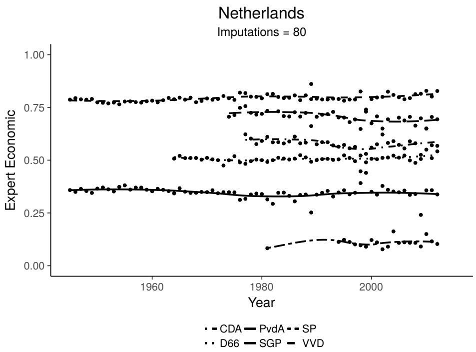
Expert	Type	Period
Castles and Mair (1984)	Expert	1982
Laver and Hunt (1992)	Expert	1989
J. Huber and Inglehart (1995)	Expert	1993
Hix and Lord (1997)	Expert	1996
Ray (1999) and Steenbergen and Marks (2007)	Expert	1988, 1992
Lubbers (2000)	Expert	2000
Warwick (2006)	Expert	2001
Benoit and Laver (2006)	Expert	2003
Kitschelt (2011)	Expert	2008
Chapel Hill Expert Survey (2015; 2017)	Expert	1999, 2002, 2006, 2010
CSES	Expert	1998, 2005, 2009, 2013
GLES	Survey	1994, 1998, 2002, 2005, 2009, 2013
Politbarometer	Survey	1987-2013
Eurobarometer	Survey	1987-1997, 1999-2000, 2002

Table 3: Overview of the data employed for Germany

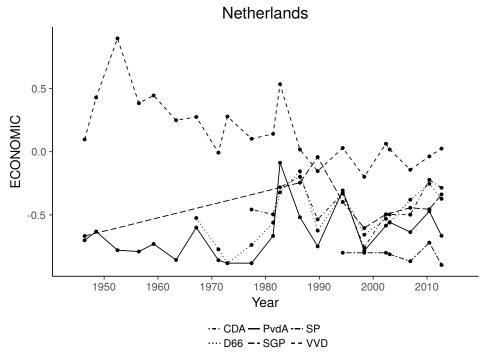
Expert	Type	Period
Castles and Mair (1984)	Expert	1984
Ray (1999) and Steenbergen and Marks (2007)	Expert	1988,1992
Laver and Hunt (1992)	Expert	1989
Hix and Lord (1997)	Expert	1996
Benoit and Laver (2006)	Expert	2003
Kitschelt (2011)	Expert	2009
Gemenis and Nezi (2012)	Expert	2011
Chapel Hill Expert Survey (2015; 2017)	Expert	1999, 2002, 2006, 2010, 2014
Eurobarometer	Survey	1980-1994, 1997, 1999-2000, 2002, 2004, 2008
EKKE	Survey	1985, 1988, 1989, 1996

Table 4: Overview of the data employed for Greece

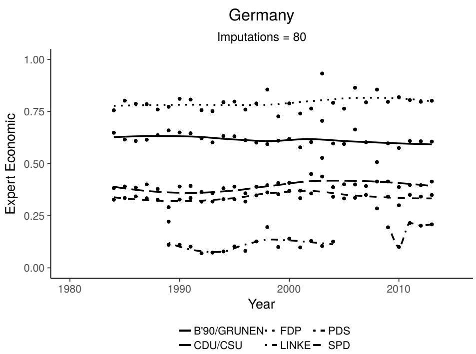
Appendix B - Economic and Social Dimensions



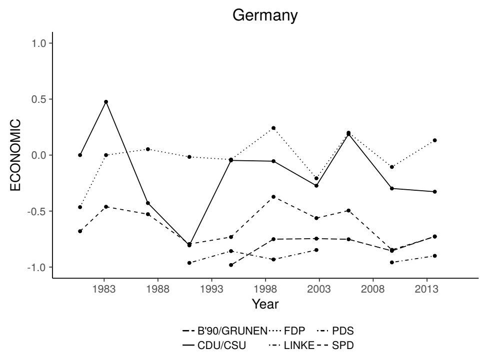
(a) Imputed Expert Positions



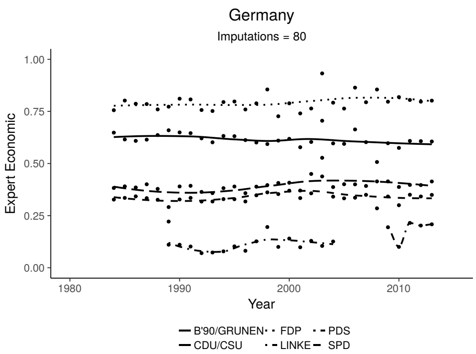
(b) Economic Positions



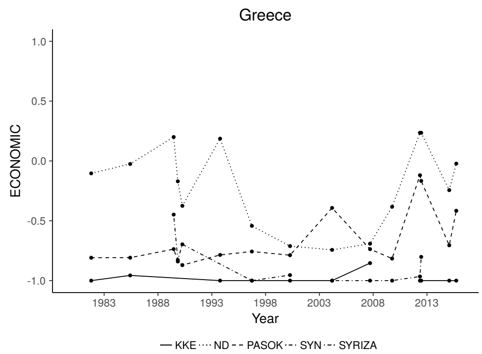
(c) Imputed Expert Positions



(d) Economic Positions

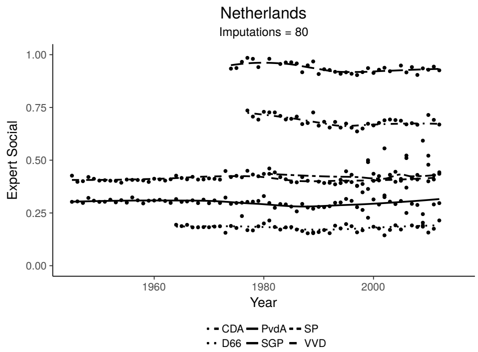


(e) Imputed Expert Positions

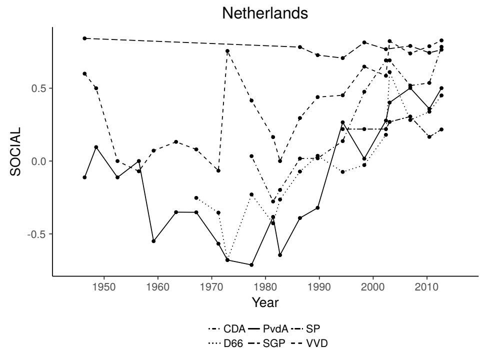


(f) Economic Positions

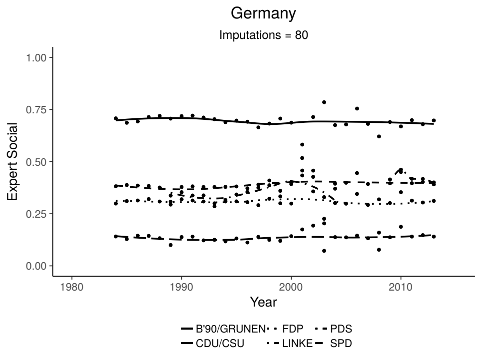
Figure 7: Comparison of imputed expert placements and the CMP Economic dimension



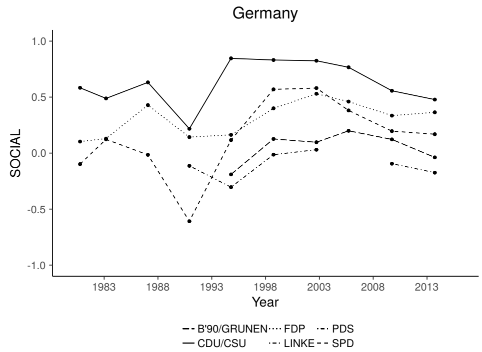
(a) Imputed Expert Positions



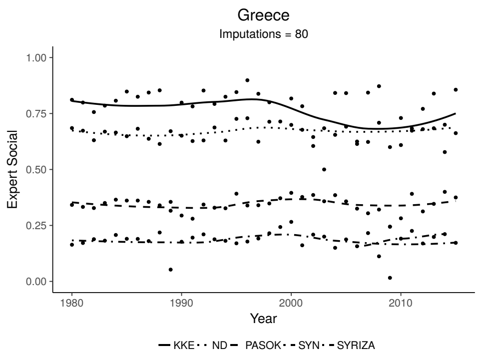
(b) Social Positions



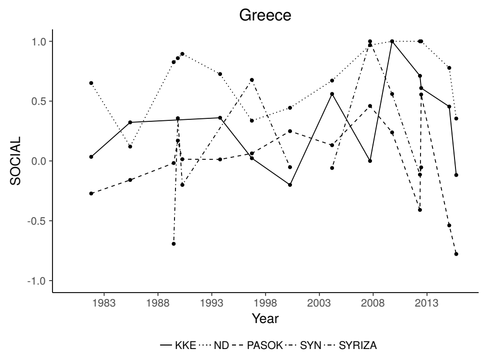
(c) Imputed Expert Positions



(d) Social Positions



(e) Imputed Expert Positions



(f) Social Positions

Figure 8: Comparison of imputed expert placements and the CMP Social dimension